mated Note

13. A Prolicion for the policion for the policion.
 2. A policion for the policion for the policion.

S. S. Howlings

14 1000 (17)

ig after at the bill Army His Control of 19009-1901 (497) by

Labrietory

HOTE OF TREESON OF

Monte a completion ex-

will be public returned, its injustion or bridged.

the content of each decrement was performed at Linear Laboratory, and a content is repeated by Massachmaette Landinte of Technology, and a property of monthly the Balliatic Massach Defense Program Office, and a property of the Balliatic Massach Defense Program Office, and a property of the Balliatic Massach Defense of the following Center under Air Furth Content F1962N-78-C-0002.

had conclosions contained in this document are those of the booking the health or be interpreted an accordantly representing the contained contained such as the United States

the majority transport that he companies and the appropried has publication.

no Recognized

was two on start Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY LINCOLN LABORATORY

THE USE OF LINEAR PREDICTION FOR THE INTERPOLATION AND EXTRAPOLATION OF MISSING DATA AND DATA GAPS PRIOR TO SPECTRAL ANALYSIS

S. B. BOWLING

Group 35

S. T. LAI

Group 37



TECHNICAL NOTE 1979-46

18 JUNE 1979

Approved for public release; distribution unlimited.

LEXINGTON

MASSACHUSETTS



Abstract

The spectral analysis of a series of equally spaced samples of a timestationary process becomes difficult when samples are missing or sizable data
gaps occur within the interval of interest. A linear prediction algorithm can
be used to fill in the missing data with estimates that are spectrally consistent with the data that are observed. Simulated and practical radar
examples demonstrate an improvement in resolution and a reduction of sidelobe
interference levels. Computer programs are provided which accomplish the
extrapolation and interpolation for complex data.

Asscust	on For	_/	4
NTIS G	All & I	V	
DOG TAR		H	
Uncamput	nced		
Justiff	laution_		
By			_
nietri	but ton/		
Trans.	artiity.	rndes	
Avail	dur TTX	3/07	
	Availat	ld/or.	
Dist	speci	al	
^	1		

CONTENTS

	PAGI
Abstract	iii
A. Problem Definition	1
B. Description of the Method	2
C. <u>Simulated Examples</u>	3
D. Radar Example	4
E. <u>Limitations of the Method</u>	5
F. Computer Codes	5
G. <u>Summary</u>	6
APPENDIX COMPUTER CODES	13
References	19

PRECEDING PAGE BLANK - NOT FILMED

v

A. Problem Definition

When a spectral transformation of a sampled process is performed, one must account for any samples that are missing. Assigning a value of zero to missing data prior to Fourier transformation, for example, introduces false frequencies and greatly increases sidelobe levels. Clearly, an interpolation scheme is needed that can cope with missing data and, at the same time, will not degrade the spectral information contained in the data that are observed.

Occasional missing samples, well separated from each other, can be estimated with simple interpolation procedures (polynomial or parabolic fits, spline fits, etc.). However, data may be missing in such quantity that conventional interpolation is inadequate; data gaps longer than the periods of the sinusoidal components in the data cannot be easily bridged with simple functions. A more sophisticated approach becomes necessary, and the use of a data-adaptive linear prediction filter is one feasible alternative.

In radar data processing, missing data or data gaps may occur for a variety of reasons:

- (a) hardware fails to transmit pulses or receive echoes properly;
- (b) radar transmits when it should be receiving echoes (range eclipsing);
- (c) resources are saturated by many targets that must be watched simultaneously (panic);
- (d) burst waveforms are purposely silent between bursts;
- (e) poor signal-tc-noise makes detections sporadically unreliable.

In any case, the missing samples (in these examples, <u>complex</u> samples with amplitude and phase) must be filled in before Doppler processing can be accomplished.

B. Description of the Method

The use of a linear prediction filter to extend a finite complex data set before Fourier transformation was first proposed and described by Bowling (1977). Applying this original algorithm, Tomlinson and Ackerson (1978) demonstrated clutter and sidelobe reduction in the Doppler processing of a train of radar pulses.

In the application of interest here, the prediction algorithm is used to predict estimates of missing data by extrapolating from observed data. For example, suppose an observation interval contains randomly missing samples and gaps. The procedure is as follows:

- (1) Locate and designate the missing samples to be estimated.
- (2) Find the longest continuous span of data within which there are no missing samples.
- (3) Calculate an N-point linear prediction filter from the longest continuous span of data found in step (2).
- (4) Calculate an estimate of each missing sample immediately to the left and to the right of the longest continuous span of data (a total of two estimates, one on each side).
- (5) Return to step (2) until all missing data have been estimated. Note that estimates from step (4) are to be treated as observations on an equal basis with the original data.

That is, the longest continuous span of data is increasing in length as estimates fill in the holes, one by one, to the left and to the right.

When the longest continous span of data finally terminates at one of the endpoints of the observation interval, estimates continue to be made toward the other endpoint until all missing points have been filled in. The length of the prediction filter may remain a constant, or vary according to the current length of the longest span of data.

C. Simulated Examples

A simple example shows the improvement in the power spectrum of a data set containing missing samples and gaps.

The real and imaginary parts of a sampled sum of three complex sinusoids are shown in Figs. 1(a,b). No noise has been added and no samples are missing. Figure 1(c) is the true power spectrum calculated with a standard FFT. No weighting function has been used.

Now if samples are randomly zeroed out and data gaps are introduced as shown in Figs. 2(a,b), the power spectrum in Fig. 2(c) shows increased side-lobe levels and false frequencies, both caused by processing without estimating the missing data.

Figures 3(a,b) show the data set after the linear prediction algorithm is applied, with the power spectrum shown in Fig. 3(c). Not only do Figs. 1(a,b) overlay with 3(a,b) almost exactly, but their respective power spectra are indistinguishable.

Another simple example demonstrates the performance of the linear prediction algorithm when data gaps occur periodically, such as is the case for a radar burst waveform.

Figures 4(a,b) represent the process of Figs. 1(a,b) for which three data gaps are present. Indeed, half of the data are missing from the observation interval, and the gaps are longer than any period exhibited in the data. The power spectrum of Fig. 4(c) is a very poor estimate of the true spectrum (Fig. 1(c)) because no gaps have been filled in. Transforming only one of the short spans of observed data gives a power spectrum with limited resolution, as shown in Fig. 4(d).

However, upon using the prediction algorithm on Figs. 4(a,b), we obtain Figs. 5(a,b) and the power spectrum in Fig. 5(c), which is an excellent estimate of the true spectrum.

In this case, the prediction algorithm has acted as a synergistic device that, by linking short pieces of data together with spectrally consistent estimates, allows a spectral transform to be performed over an effectively longer piece of data. The whole, then, has more resolving power than any of its parts.

It should be pointed out that the data gaps need not be periodic or equal in length in order for the prediction algorithm to fill them in.

D. Radar Example

Radar is often used to identify targets from the time history of the velocity spectrum of the target's motion about its center of mass. A series or burst of radar pulses is Fourier analyzed, and the target's velocity

spectrum is observed. If not accounted for in the processing, missing pulses can introduce false velocity components and lead to an incorrect characterization of the target.

For example, Fig. 6(a) shows the evolution of the velocity spectrum of a tumbling object for which missing data and data gaps exist and are set to zero in the radar pulse train. No estimation for the missing pulses has been done. It is therefore not clear if the velocities indicated are actually from the target or are an artifact of the missing data. Figure 6(b) shows the evolution of the same velocity spectrum upon using the prediction algorithm before Fourier transformation. The disappearance of some of the velocities cleans up the spectral history and indicates which velocity components actually characterize the target.

E. Limitations of the Method

Implicit in the use of a linear prediction filter is the assumption that the data from which the filter is derived are coherent. The process being sampled must be approximately stationary during the observation interval which is being analyzed and within which the missing data and data gaps may occur.

Also, the prediction filter works best when the spectral components are approximately pure tones, confined to locally narrow bandwidths spaced within the Nyquist bounds of the spectral transform domain.

F. Computer Codes

Appendix I lists three self-contained subroutines that accomplish the method described in Section B. GAPFIL, the driving subroutine, calls COEFF

and LNPRED, which calculate prediction filter coefficients and perform linear predictions, respectively. Although the code is designed for complex data, purely real data can be treated by setting the imaginary components to zero before calling GAPFIL. See Bowling (1977) for a more detailed description of the linear prediction procedure.

G. Summary

This paper proposes the use of a linear prediction algorithm to fill in missing data and data gaps that may occur within an observation interval over which a spectral transform is to be made. False frequencies and sidelobe interference, which are artifacts of the missing samples, can be surpressed or eliminated by replacing the missing samples with estimates that are spectrally consistent with neighboring observed data. Large gaps can be smoothly bridged that otherwise could not be satisfactorily interpolated by simpler schemes.

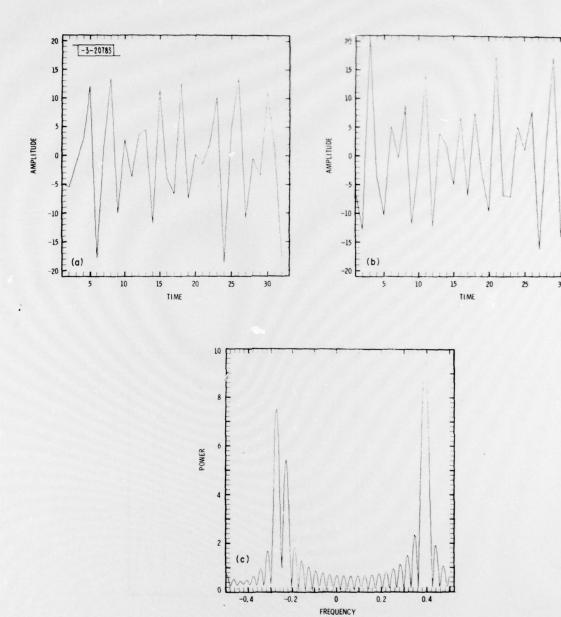


Fig. 1. (a) Real part of the sum of three complex sinusoids; no samples are missing. (b) Imaginary part of the sum of three complex sinusoids; no samples are missing. (c) Power spectrum of 1(a,b).

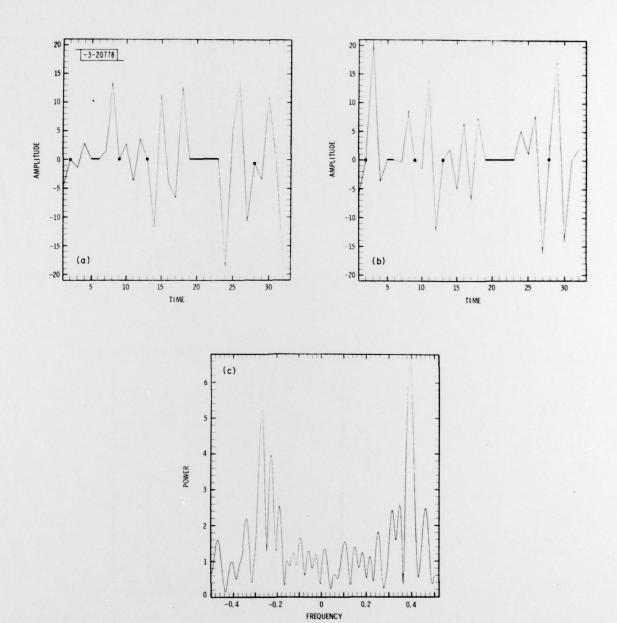


Fig. 2. (a) Real part of Fig. 1 with randomly missing data and data gaps. (b) Imaginary part of Fig. 1 with randomly missing data and data gaps. (c) Power spectrum of 2(a,b).

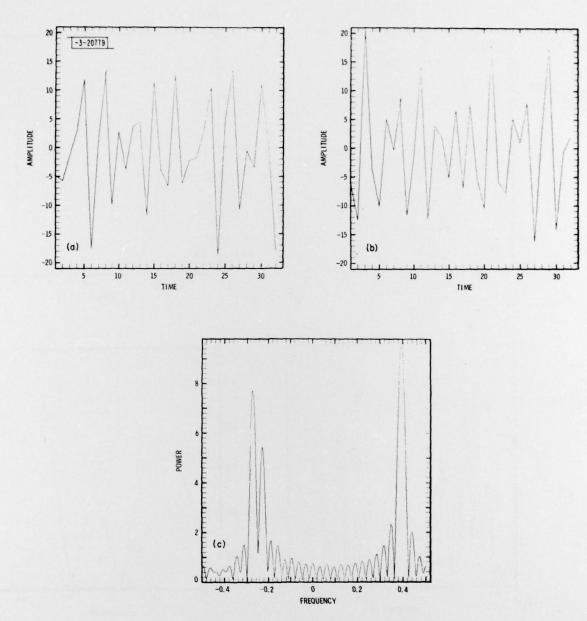


Fig. 3. (a) Real part of Fig. 2 after application of the linear prediction algorithm. (b) Imaginary part of Fig. 2 after application of the prediction algorithm. (c) Power spectrum of 3(a,b).

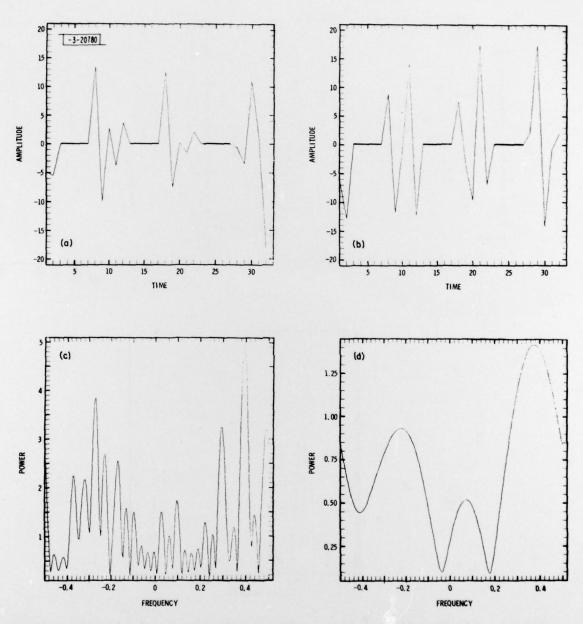


Fig. 4. (a) Real part of Fig. 1 with data gaps. (b) Imaginary part of Fig. 1 with data gaps. (c) Power spectrum of 4(a,b). (d) Power spectrum of the central short data span in 4(a,b).

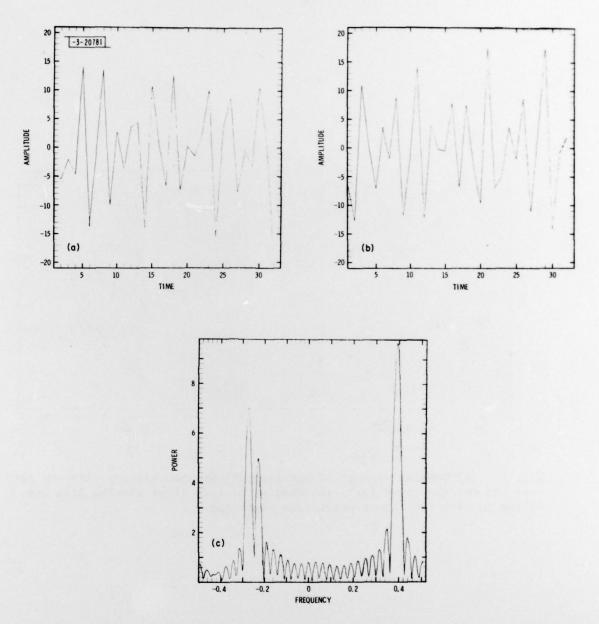


Fig. 5. (a) Real part of Fig. 4 after gaps are filled in with the linear prediction algorithm. (b) Imaginary part of Fig. 4 after gaps are filled in with the prediction algorithm. (c) Power spectrum of 5(a,b).

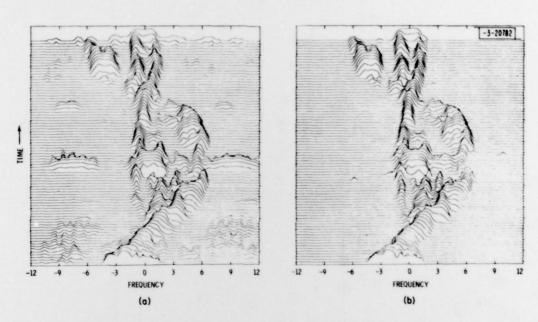


Fig. 6. (a) Doppler history of tumbling object when missing data and data gaps are not accounted for. (b) Doppler history after missing data are filled in with the linear prediction algorithm.

APPENDIX COMPUTER CODES

Listings for subroutines:

- 1) GAPFIL driving program
- 2) COEFF filter coefficient calculations
- 3) LNPRED estimation by linear prediction

SUBBOUTINE GAPPIL (XREAL, XIMAG, NPTS, X, A, PM, AA, B1, B2)

C

COMPLEX X, A, AA, B1, B2, ARRAY DIMENSION X(1), A(1), XREAL(1), XIMAG(1) DIMENSION PM(1), AA(1), B1(1), B2(1) DIMENSION ARRAY(256), MISS(128)

0000

INPUTS ARE: KREAL, XIMAG, NPTS

OUTPUTS ARE XREAL, XIMAG, X, A

0000

C

C

c c

C

C

C

C

C

C

C

THE ARGUMENTS OF THIS SUBROUTINE ARE:

XREAL (NPTS) = ARRAY CONTAINING THE REAL PART OF THE DATA
XIMAG (NPTS) = ARRAY CONTAINING THE IMAGINARY PART OF THE DATA
NPTS= THE NUMBER OF DATA POINTS (INCLUDING MISSING POINTS)
X (NPTS) = COMPLEX ARRAY OF DATA POINTS USED IN PREDICTION
A (NCOEPP) = COMPLEX ARRAY OF PREDICTION COEPPICIENTS GENERATED
WITHIN THIS SUBROUTINE (MCOEPP WILL NOT EXCEED NPTS/4)
PM (NCOEPF) = ARRAY OF ERROR POWER HISTORY AS PILTER IS BUILT
AA (NCOEFF), B1 (NPTS), B2 (NPTS) = WORK ARRAYS

IN THE ABOVE DESCRIPTION, THE MINIMUM SIZES OF THE ARRAYS AS USED IN THIS SUBROUTINE ARE SPECIFIED

0000

C

C

C

C

C

THE LONGEST SPAN OF CONTIGUOUS DATA IS FIRST LOCATED WITHIN THE INTERVAL NPTS. THEN THE MISSING DATA IN THE ADJACENT SAMPLE POSITIONS AFE PREDICTED USING A LINEAR PREDICTION ALGORITHM. AS MISSING POINTS AND GAPS ARE PILLED IN, THE BASIS FROM WHICH THE PREDICTIONS ARE BEING MADE INCREASES IN LENGTH; NCOEPP THE LENGTH OF THE PREDICTION FILTER, IS NOT ALLOWED TO EXCEED NPTS/4. THE DATA ARRAYS ARE RETURNED WITH THE GAPS AND MISSING POINTS FILLED IN.

0000

CC

SEE LINCOLN REPORT RMP-122 FOR DESCRIPTION OF LINEAR PREDICTION

IF NPTS IS GREATER THAN 256 , INCREASE DIMENSION OF 'ARRAY' TO ACCOMODATE
IF MORE THAN 128 POINTS ARE MISSING, INCREASE DIMENSION OF 'MISS'.

0000

C

C

STEP 1 : IDENTIFY MISSING DATA.

MISSING DATA APE THOSE CONSIDERED TO HAVE VALUES OF ZERO. IF THIS CRITERION IS NOT APPROPRIATE FOR YOUR APPLICATION, THEN YOU WILL HAVE TO REWRITE STEP 1. NOMISS IS THE NUMBER OF MISSING DATA POINTS, AND MISS(J) IS THE LOCATION OF THE JTH MISSING POINT. THUS, MISS(J) CAN HAVE ANY VALUE BETWEEN 1 AND MPTS.

```
C
      I = 0
      J = 0
    9 I = I+1
      X(I) = CAPLI(IRBAL(I), XIRAG(I))
      IF (IREAL (I) . BQ. 0. 0. AND. XIMAG (I) . EQ. 0. 0) J=J+1
      IF (IRBAL (I) . EQ. 0.0. AND . XIMAG (I) . EQ. 0.0) MISS (J) = I
      IP (I.LT. MPTS) GO TO 9
      NOMISS = J
      IF (WOMISS. EQ. 0) RETURN
C
C-
C
      STEP 2 : SEARCH FOR LONGEST SPAN OF CONTINUOUS DATA,
C
C
      AND GENERATE 2 NUMBERS (L1, L2) CHARACTERIZING THE SPAN.
C
      L1 = 0
      L2 = MISS (1)
      H1 = 1
      M2 = 1
      MAX= MISS(1)
      IF (NOMISS.EQ. 1) GO TO 30
      DO 20 J = 2, NOMISS
JDIF = MISS(J) - MISS(J-1)
      IF (JDIF.LE.MAX) GO TO 20
      M1 = J-1
      M2 = J
      L1=MISS(M1)
      L2=MISS(M2)
      MAX= JDIP
   20 CONTINUE
   30 JDIF = NPTS-MISS (NOMISS)
      IF (JDIF. LE. MAX) GO TO 34
      M1 = NOMISS
      M2 = NOMISS+1
      L1 = MISS (M1)
      L2 = NPTS + 1
C
C-
C
C
      STEP 3 : DEPINE NUMBER OF LINEAR PREDICTION COEFFICIENTS,
C
                 AND PERFORM PREDICTION.
C
   34 NCOEFF = (L2-L1)/2
      IF (MCOBPP.LT.1) NCOEFF=2
      IF (NCOEFF.GT. (NPTS/4)) NCOEFF= (NPTS/4)
I1 = L1+1
      12 = L2-1
      J = 0
      DO 40 I = 11, 12
      J = J+1
   40 ARRAY(J) = X(I)
      MP = J
      MPP2 = J+2
C
```

```
CALL COSFF (MP, ARRAY, MCORFF, A, PM, PO, AA, B1, B2)
C
          CALL LEPRED (NP, HPP2, ARRAY, MCORFF, A)
C
C
C-
C
          STEP 4 : RE-ESTABLISH REAL AND IMAGINARY ARRAYS.
C
         IF (L1.GT.0) XREAL(L1) = REAL(ARRAY(1))
IF (L1.GT.0) XIMAG(L1) = AIMAG(ARRAY(1))
IP (L1.GT.0) X(L1) = ARRAY(1)
C
          IP (L2.LE.MPTS) IRBAL(L2) = BBAL(ARRAY(MPP2))
IP (L2.LE.MPTS) IIMAG(L2) = AIMAG(ARRAY(MPP2))
IP (L2.LE.MPTS) I(L2) = ARRAY(MPP2)
C
          H1 = H1-1
          IP (M1.GE. 1) L1=MISS(M1)
         IP (M1.LT.1) L1=0
M2 = M2+1
         IP (M2.LE.MONISS) L2=MISS(M2)
IP (M2.GT.MONISS) L2=MPTS+1
IP (M1.LT.1.AND.M2.GT.MONISS) GO TO 50
          GO TO 34
     50 RETURN
          END
```

```
SUBROUTINE COEFF (NPTS, X, NCOEFF, A, PN, PO, AA, B1, B2)
C
   THIS SUBROUTINE CALCULATES THE COMPLEX BURG COEFFICIENTS
C
   (HERE CALLED AFRAY A). THE ALGORITHMS USED HERE ARE AN EXTENSION OF
C
   ALGORITHMS DESCRIBED BY ANDERSON (GEOPHYSICS, VOL 39, PEB. 1974)
C
   TO THE CASE OF A COMPLEX SERIES.
C
C
   PROGRAMMED BY S.B. BOWLING, MIT-LINCOLN LABORATORY, SEPT. 1976.
      COMPLEX X, A, AA, B1, B2, XNOM, DEN, TWO
      DIMENSION X(1), A(1), AA(1), B1(1), B2(1), PR(1)
      TWO=CMPLX (2.0,0.0)
       PO=0.0
      DO 10 IT= 1, NPTS
      DUMMY= X(IT) *CONJG (X(IT))
  10 PO= PO+ DUMMY
      PO= PO/PLOAT (NPTS)
      NM1=NPTS-1
      B1(1)=X(1)
      B2 (NM1) = X (NPTS)
      DO 20 IT=2, NM1
      B1 (IT) = X (IT)
      ITM1=IT-1
      B2(ITM1) = X(IT)
      DO 50 M=1, NCOEFF
      MM 1=8-1
      NMM=NPTS-M
      IF (M . EQ. 1) GO TO 25
      DO 21 IT=1,MM1
  21 AA (IT) = A (IT)
       DO 22 IT=1, NMM
       B1(IT) = B1(IT) - CONJG(AA(MN1)) *B2(IT)
  22
      B2 (IT) = B2 (IT+1) - AA (MM1) + B1 (IT+1)
      XNOM=CMPLX (0.0,0.0)
       DEN=CMPLX (0.0,0.0)
      DO 30 IT=1, NAM
       XNOM=XNOM + B2(IT) *CONJG(B1(IT))
      DEN=DEN + B1 (IT) *CONJG (B1 (IT) ) + B2 (IT) *CONJG (B2 (IT))
      IF ( REAL (DEN) . EQ. 0.0) GO TO 35
       A(M) = TWO+ (XNOM/DEN)
      GO TO 36
      A(M) = CMPLX (0.0,0.0)
  36 POWER=PO
       IF (N .GT. 1) POWER=PM (N-1)
       DUMMY= A (M) *CONJG (A (M))
       PH (H) = POWER* (1.0 - DUMMY)
       IF (H . EQ. 1) GO TO 50
      DO 40 IT=1,881
      A(IT) = AA(IT) - A(H) + CONJG(AA(H-IT))
  50
      CONTINUE
       RETURN
       END
```

CONT.

经验证金额

```
SUBBOUTINE LNPRED (N1, N2, X, NCOBPP, A)
   THIS SUBROUTINE LINEARLY EXTENDS THE COMPLEX DATA ARRAY "X"
C
   FROM N1 POINTS TO N2 POINTS. THE ORIGINAL DATA IS CENTERED
   IN AN ARRAY OF W2 ELEMENTS AND BOTH A FORWARD AND BACKWARD
   EXTENSION ARE PERFORMED UNTIL THE TOTAL BURBER OF POINTS IS N2.
   NI= ORIGINAL NUMBER OF POINTS IN ARRAY X
   N2=NUMBER OF POINTS TO WHICH ARRAY X IS EXTENDED
C
   X = COMPLEX ARRAY OF DATA SAMPLES, DIMENSION AT LEAST #2
   NCOEFF= NUMBER OF PREDICTION FILTER COEFFICIENTS
   A= COMPLEX ARRAY OF FILTER COEPFICIENTS, DIMENSION NCOEPF
C
   PROGRAMMED BY S.B. BOWLING, MIT-LINCOLN LABORATORY, SEPT. 1976.
C
      COMPLEX X, A
      DIMENSION X(1), A(1)
   SET UP LIMITS FOR DO LOOPS
      L1= N2/2 - N1/2
      L2= N2/2 + N1/2
      IF ( NOD (N1,2) . EQ. 1 ) L2=L2+1
C
   SHIPT ORIGINAL DATA TO MIDDLE OF ARRAY X
C
      DO 100 I=1, #1
      J = N1 - (I-1)

K = L2 - (I-1)
 100
     X(K) = X(J)
   DO FORWARD PREDICTION
C
      N3 = N2 - L2
      DO 200 I=1.N3
      J = L2 + I
      X(J) = CMPLX(0.0,0.0)
      DO 200 K=1,NCOEFF
     X(J) = X(J) + A(K) *X(J-K)
C
   DO BACKWARD PREDICTION
C
      DO 300 I=1,L1
      J = L1 - (I - 1)
      X(J) = CMPLX(0.0,0.0)
      DO 300 K=1, NCOEFF
 300
      X(J) = X(J) + CONJG(A(K)) * X(J+K)
C
      RETURN
      END
```

000

References

- Bowling, S. B., "Linear Prediction and Maximum Entropy Spectral Analysis for Radar Applications," Project Report RMP-122, Lincoln Laboratory, M.I.T. (24 May 1977), DDC-AD-AO4817/7.
- Tomlinson, P. G., and G. A. Ackerson, "Air Vehicle Detection Using Advanced Spectral Techniques," Proceedings of the RADC Spectrum Estimation Workshop, Rome Air Development Center, Rome, New York (May 1978).

THE DAILS SEED

UNCLASSIFIED SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) READ INSTRUCTIONS REPORT DOCUMENTATION PAGE BEFORE COMPLETING FORM 2. GOVT ACCESSION NO. 3. RECIPIENT'S CATALOG NUMBER ESD-TR-79-160 TYPE OF REPORT & PERIOD COVERED TITLE (and Subtitle) The Use of Linear Prediction for the Interpolation and Technical Note Extrapolation of Missing Data and Data Gaps Prior to Spectral Analysis . PERFORMING ORG. REPORT NUMBER Technical Note 1979-46 7. AUTHORIS CONTRACT OR GRANT NUMBER(s) Stephen B. Bowling and Shu T. Lai F19628-78-C-0002 10. PROGRAM ELEMENT, PROJECT, TASK AREA WORK UNIT NUMBERS PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Project No. 8X363304D215 Lexington, MA 02173 12. REPORT DATE 11. CONTROLLING OFFICE NAME AND ADDRESS
Ballistic Missile Defense Program Office 18 June 1979 Department of the Army 5001 Eisenhower Avenue NUMBER OF PAGES Alexandria, VA 22333 26 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 15. SECURITY CLASS. (of this report) Electronic Systems Division Unclassified Hanscom AFB 15a. DECLASSIFICATION DOWNGRADING SCHEDULE Bedford, MA 01731 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) 18. SUPPLEMENTARY NOTES None 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) radar signal analysis predictive deconvolution ARMA linear prediction filter data adaptive maximum entropy interpolation and extrapolation autoregressive decomposition spectral analysis 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The spectral analysis of a series of equally spaced samples of a time-stationary process becomes difficult when samples are missing or sizable data gaps occur within the interval of interest. A linear prediction algorithm can be used to fill in the missing data with estimates that are spectrally consistent with the data that are observed. Simulated and practical radar examples demonstrate an improvement in resolution and a reduction of sidelobe interference levels. Computer programs are provided which accomplish the extrapolation and interpolation for complex data.

DD FORM 1473 EDITION OF I NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

207650

Du